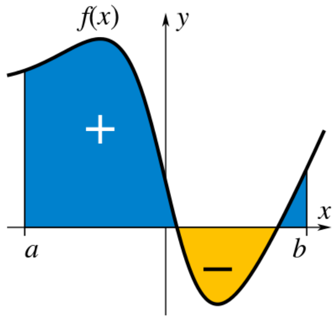
How to calculate integrals

Given a function f of a real variable x and an interval [a, b] of x-axis, the *definite integral*

$$\int_a^b f(x)dx)$$

is defined as the signed area of the region in the xy-plane that is bounded by the graph of f, the x-axis and the vertical lines x=a and x=b with a<b. The area above the x-axis adds to the total and that below the x-axis subtracts from the total. **[1]**



The fundamental theorem of calculus connects differentiation with the definite integral:

$$\int_a^b f(x)dx = F(b) - F(a)$$

with F being the *antiderivative* (indefinite integral) and its property F ' (x) = f (x).

1. Integral of a Polynomial

A function f is called *polynomial* if it has the structure

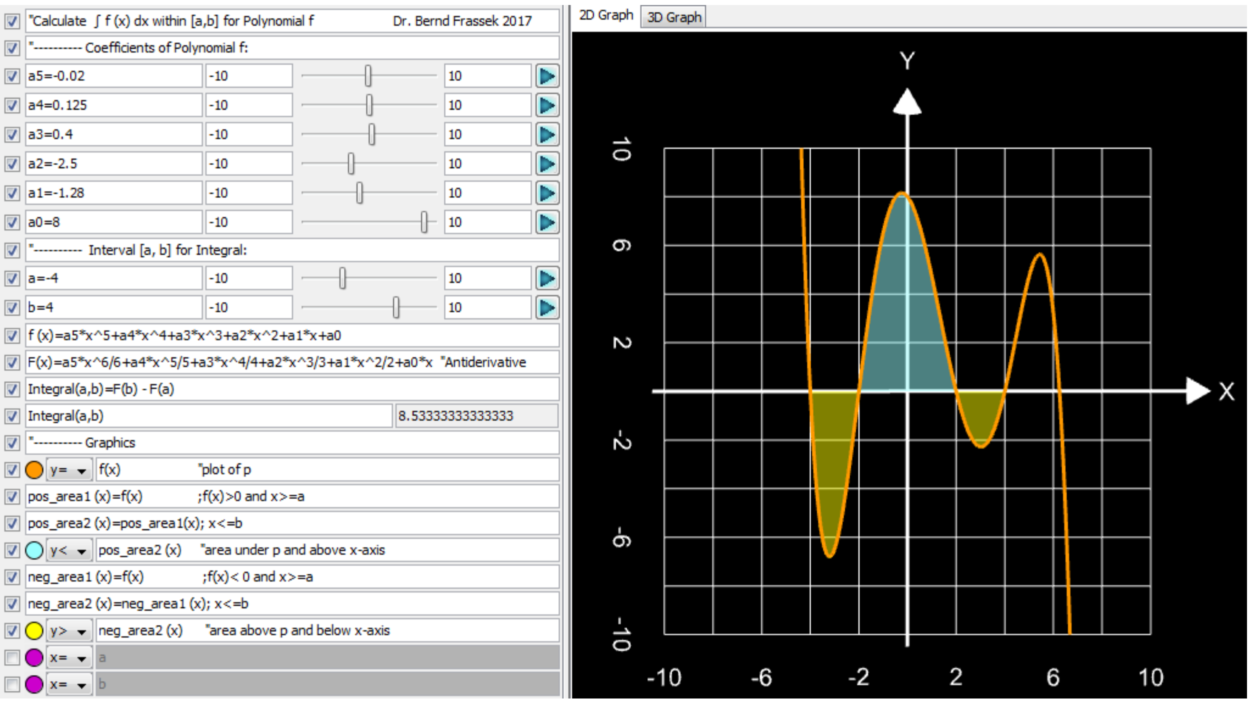
$$f(x) = a_n\, x^n + \cdots + a_2\, x^2 + a_1\, x + a_0$$

With a_0, …, a_n being a real number and all exponents being natural.

Its anti-derivative is

$$F(x) = a_n \frac{1}{n+1} x^{n+1} + \cdots + a_2 \frac{1}{3} x^3 + a_1 \frac{1}{2} x^2 + a_0 x$$

The GC3 file **Integral of Polynomial.gc3** calculates the definite integral of f in the interval [a, b].



The coefficients and the interval can be set with sliders. The function is plotted in orange, areas above the x-axis are filled with blue, areas below the x-axis are filled with yellow.

Programming Details:

- For filling the area below the graph of f and above the x-axis the plot is done with an inequation "y<" fulfilling the condition

  f(x)>0 AND x≥a AND x≤b

  Since GC3 can 'only' handle two expressions combined with AND / OR, there are two lines needed to implement this condition:

  pos_area1(x)=f(x)   ;f(x)>0 and x>=a
  pos_area2(x)=pos_area1(x) ; x<=b

- **Important for all files/examples**: make sure that you **always set a lower than b**; otherwise the areas are not colored properly resp. additional case differentiations have to be done.

If the *absolute area A* below the graph of f within the interval [a, b] should be calculated, the interval has to be divided in pieces $[a,x_1]$, $[x_1,x_2]$, …, $[x_n,b]$ with $x_1 \ldots x_n$ being the roots of f.

The absolute area A is then calculated piecewise:

$$A = \left|\int_a^{x_1} f(x)dx\right| + \left|\int_{x_1}^{x_1} f(x)dx\right| + \cdots \left|\int_{x_n}^{b} f(x)dx\right|$$

In order to avoid the determination of the roots the integral of | f(x) | can be calculated by means of Numerical Integration (see below).

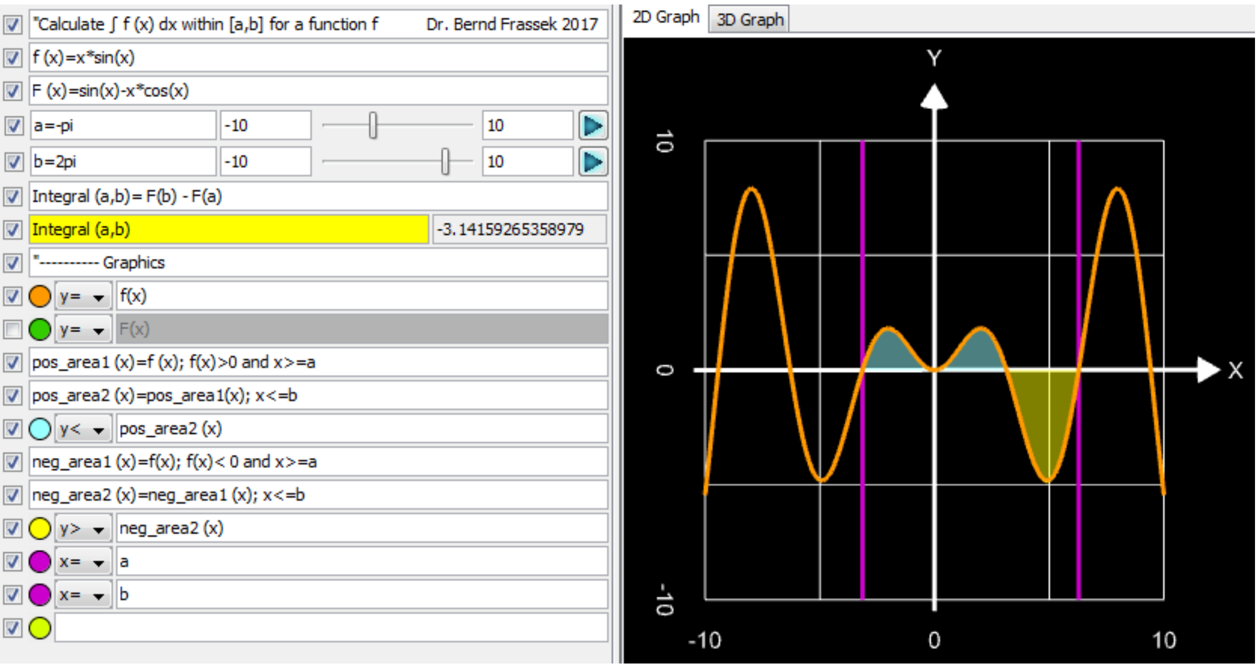
2. Integral of Functions with known Antiderivatives

If the antiderivative of a function exists and is known **[2]**, the file of the last section can be simplified (**Integral with known Antiderivative.gc3**).

The function has to be input in 'f(x) = …', e.g. f(x) = x sin(x). The antiderivative of f must be input in 'F(x)=…' which is sin(x) – x cos(x) for this example.

The interval [a, b] can be set with sliders.

GC3 calculates the integral which – π for the example.



Antiderivatives can be found in literature and in the Internet (see references at the end).

3. Numerical Integration

Another way for solving integrals is *numerical integration*. Especially when the antiderivative is hard to find or even does not exist (see below) numerical integration is the only way to solve the integral.

There are many methods for numerical integration. A quick and pretty accurate method to be used with GC3 is the *Composite Simpson Rule* [3], [4].

Suppose that the interval [a, b] is split up into 2n intervals. Then, the composite Simpson's Rule is given by

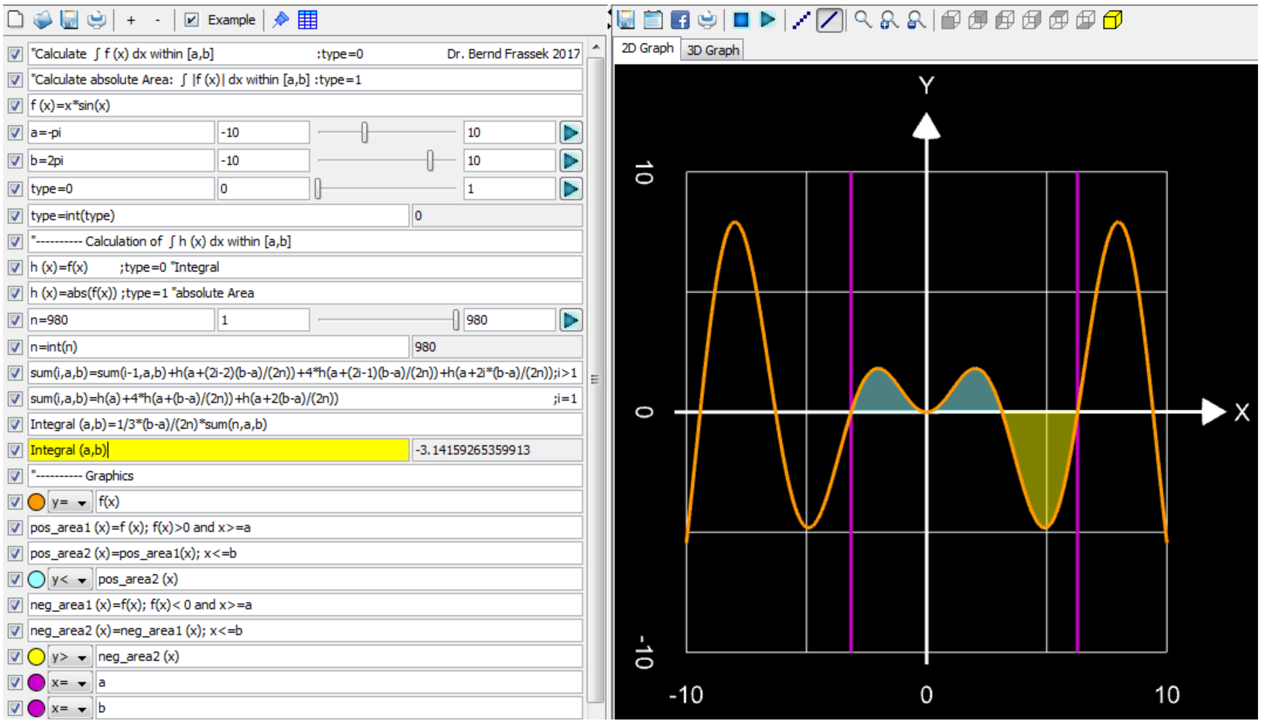
$$\int_a^b f(x)dx \approx \frac{1}{3}\frac{b-a}{2n}\sum_{i=1}^{n}[f(x_{2i-2}) + 4\,f(x_{2i-1}) + f(x_{2i})]$$

with

$$x_i = a + i \cdot \frac{b-a}{2n} \qquad \text{for} \quad i = 0, 1, \dots, 2n\text{-}1$$

Note:
All of the following examples were calculated with an older version of GC3 with n=980 being the maximum possible number for a recursion. With the actual version there is (theoretically) no limitation for n.

The above formulas are is implemented in the file **Integral & Absolute Area.gc3**:



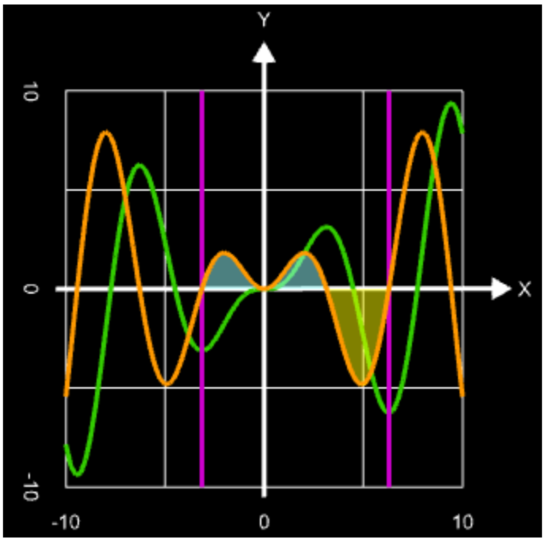
Programming Details:

- The sum is calculated with a recursion over i with i = 1 to n.
- In order to study the influence of n on the accuracy, n can be set with a slider. In the current version the number of recursions is limited; therefore, the upper border for n has to be tried out.
- In fact, the integrated function is h(x). This allows to calculate both, the integral and absolute area A by setting the value of the variable *type* with a slider:
  - If type = 0, the integral is calculated with h(x) = f(x).
  - If type = 1, the absolute area A is calculated with h(x) = |f(x)|.

  There is no need to divide the interval [a, b] in pieces depending on the roots of f – GC3 does that for you ☺.

- If you want to plot the graph of the antiderivative (integral function), you can add a further line using a regular function

in the graphics section of the file, resulting in this:



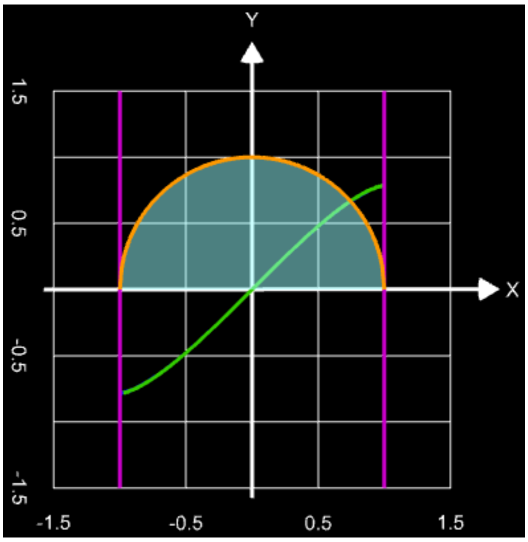
Accuracy

For the example of the last section f(x) = x sin(x) with n = 980 and GC3's standard math library 'Double' the numerically calculated value for the integral is -3.14159265359913 which compared with the exact value of -π results in an absolute error of 9.33 E-13.

If the slope of the function at the borders of the interval is ± ∞, the error will become significantly larger.

This, e.g. may happen with "elliptical' functions containing square roots, such as

$$\int_{-1}^{1} \sqrt{1-x^2}dx$$



The green plot shows the numerically calculated antiderivative F(x) = 0.5 (asin (x) + x • f(x)).

The calculated result is 1.57078884167346 which compared with the exact value of 0.5 π results in an absolute error of 7.49 E-6 (standard math library 'Double').

A further investigation of the error committed by the Composite Simpson Rule is not subject of this paper.

Further Examples

- Solve the integral

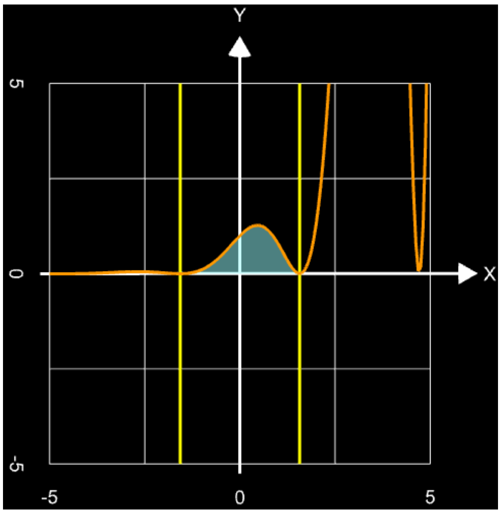
$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{x} \cos{(x)}^2 \; dx$$

Can also be solved in closed form with the antiderivative F of of

$$F(x) = e^{x} \frac{2\sin(2x) + \cos(2x) - 5}{10}$$

Absolute error: 1.2 E-12.

File: **Integral e^x cos(x)^2.gc3**



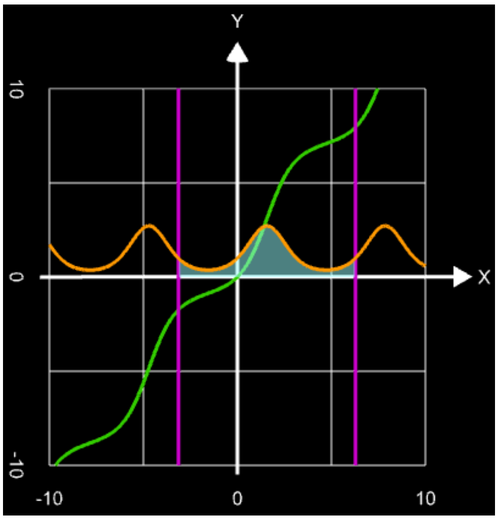
- Solve the integral

$$\int_{-\pi}^{\pi} e^{\sin(x)} \, dx$$

There is no antiderivative in closed form.

Absolute error: 1 E -14

File: **Integral e^sin(x) dx.gc3**



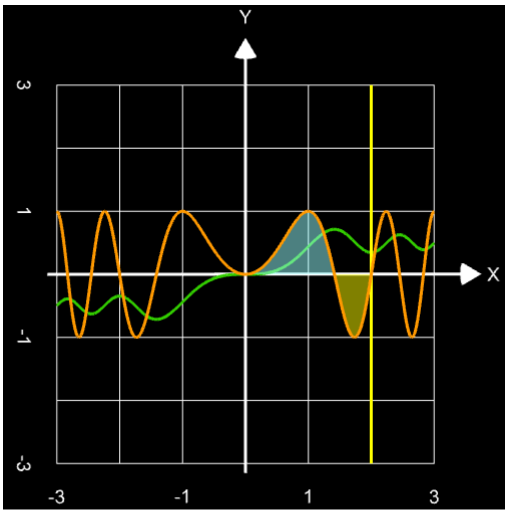
- Fresnel integral S(x) **[5]**

This is used in optics and defined as

$$S(x) = \int_{0}^{x} sin\left(\frac{\pi}{2} t^2\right) \, dt$$

For x=2 the result is 0.34341567836221 with an absolute error of 1.5 E -12

File: **Fresnel Integral S(x).gc3**



- **Normal Distribution** (file **Normal Distribution.gc3**)

The **Normal** (or **Gaussian**) **Distribution** (also known as "bell curve") is a very common continuous probability distribution for measurement values. The *probability density* is given by

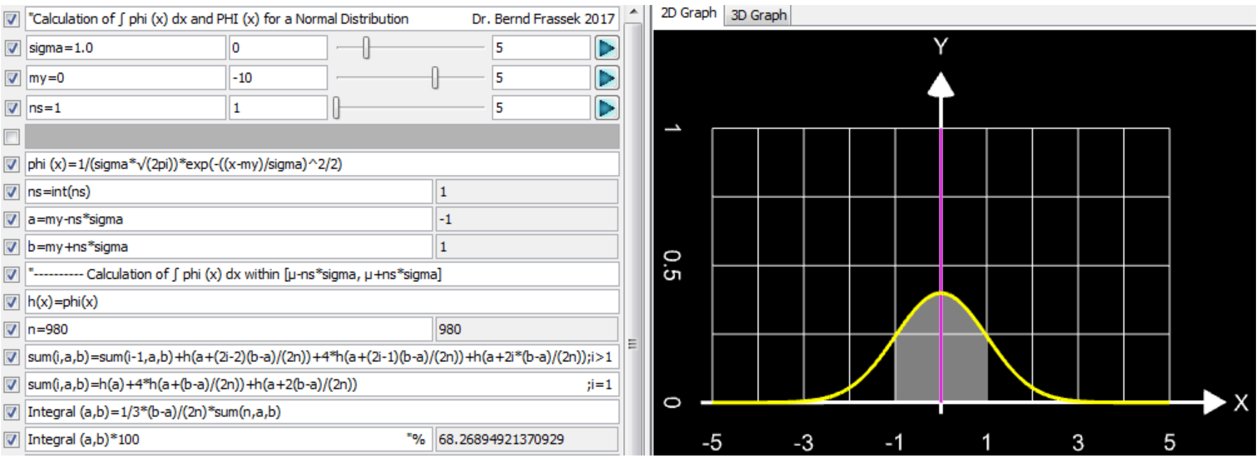
$$\varphi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where μ is the *mean value* (expectation value) and σ² is the *variance.* **[6]**

To answer the typical question 'how many of the measured values are within the range μ ± multiples of sigma?" (white area in the plot), the integral

$$\int_a^b \varphi(x)dx$$

with a= μ-ns•σ and b= μ+ns•σ has to be calculated. For this integral there exists no antiderivative in closed form; so it must be calculated numerically:



For [μ-σ, μ+σ] i.e. ns=1, the result is ~ 68.3 % whereas for .[μ-2σ, μ+2σ] i.e. ns=2 the result is ~ 95,4 %.

An extension of the task is plotting the *Cumulative Distribution Function* (CDF) [7] given by

$$\Phi(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

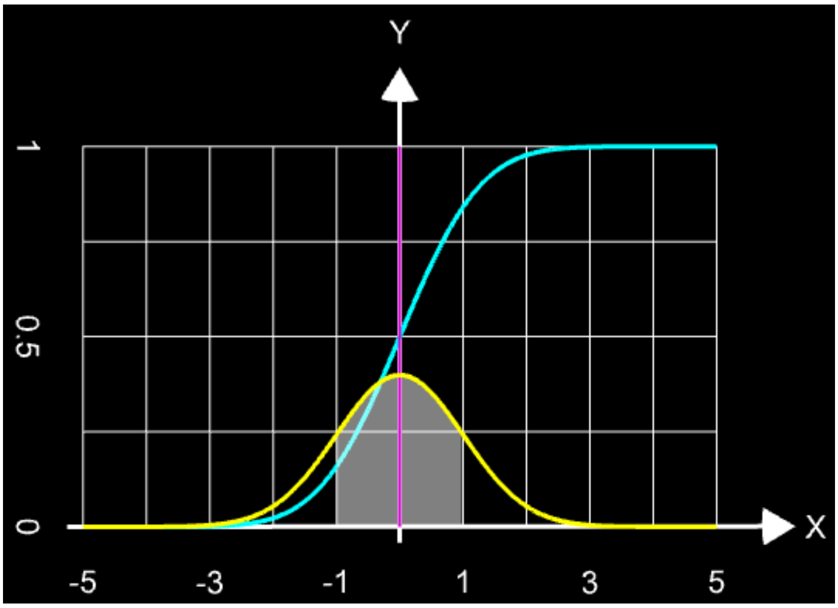
which can be re-written as

$$\Phi(\mathrm{x}) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$$

with erf (x) being the general Error Function **[8]**:

$$\mathrm{erf}\ (x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

```
"---------- Calculation of  Cumulative Distribution Function PHI (x) with Error Function erf(x)
f (t)=2/√pi*exp(-t^2)
sum2 (i,b)=sum2(i-1,b)+f((2i-2)b/(2n))+4*f((2i-1)b/(2n))+f(2i*b/(2n));i>1
sum2 (i,b)=f(0)+4*f(b/(2n))+f(2b/(2n));i=1
erf (b)=1/3*b/(2n)*sum2(n,b) "general error function
PHI(x)=0.5 (1+erf ((x-my)/(sigma√2)))
"---------- Graphics
y=  phi (x)
x=  my
area (x)=phi (x); x>=a and x<=b
y<  area (x)
y=  PHI(x)
```



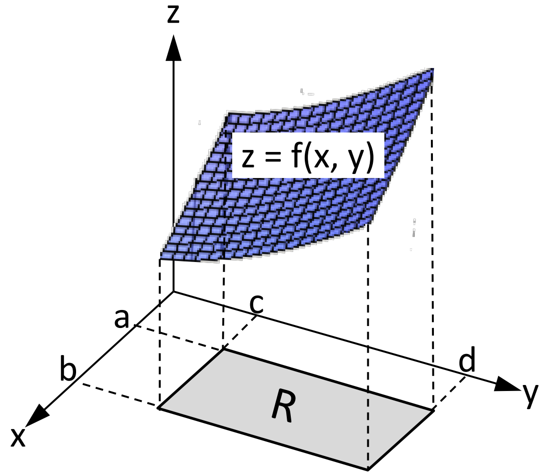
Note:

- For calculating erf (x), the parameter a is omitted in the Simpson Rule since a = 0.
- **Important**: Because the plot contains two integrals which are calculated recursively, the sum in the recursion for erf (x) has to be renamed ('sum2').

4. Double Integrals (Surface Integrals)

Integrals of a function of two variables f(x, y) over a region R in the xy-plane are called *double integrals*.

While the definite integral of a positive function of one variable f (x) represents the **area** of the region between the graph of the function and the x-axis, the double integral of a positive function of two variables f(x, y) represents the **volume** of the region between the surface defined by the function (on the three-dimensional Cartesian plane where z = f(x, y), blue area in drawing) and the region R in the xy-plane (grey area in drawing). [9]



For two basic types of functions it is shown how you plot and calculate the volume for functions f(x,y) with GC3, using numerical integration.

Type 1: $f(x, y) = f_x(x) \cdot f_y(y) + C$

The function f(x, y) consists of a term f_x having only the variable x, multiplied by a term fy having only the variable y, and a constant C.

Examples:	f(x, y)= sin(x) cos(y)	$f_x(x)$= sin(x)	$f_y(y)$=cos(y)	C= 0
	f(x, y) = √(1-y^2) - 0.5	$f_x(x)$= **1**	$f_y(y)$= √(1-y^2)	C= - 0.5
	f(x, y)= x sin(y) $e^{(-x)}$+1	$f_x(x)$= x $e^{(-x)}$	$f_y(y)$=sin(y)	C= 1
	f(x, y) = sin(x) √(1-y^2) x+2	$f_x(x)$= x sin(x)	$f_y(y)$= √(1-y^2)	C= 2

For this type of function the integral

$$\int_a^b \int_c^d f(x,y)\,dy\,dx$$

can be calculated in the following way:

$$= \int_a^b \int_c^d \left[f_x(x) \cdot f_y(y) + C\right] dy\,dx$$

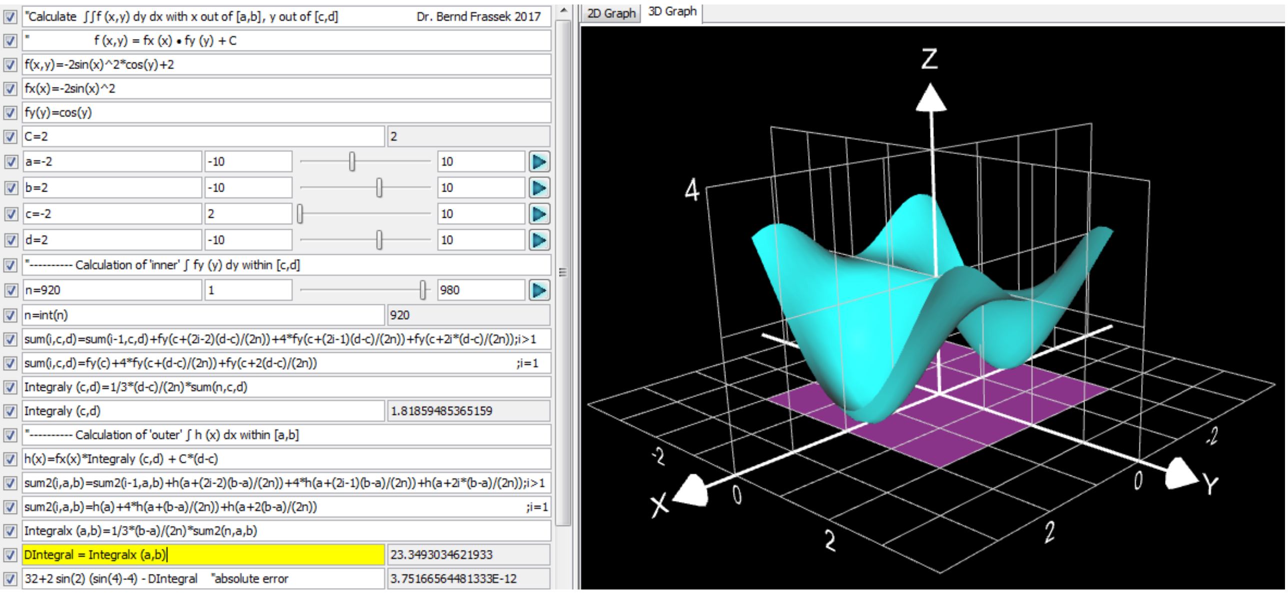
$$= \int_a^b \left[f_x(x) \cdot \int_c^d f_y(y)\,dy + C(d-c) \right] dx = \int_a^b h(x)\,dx$$

i.e. $f_y(y)$ has to be integrated at first (blue integral), followed by an integration of h(x) (red integral).

The above formulas are implemented in the file **DInt [fx(x) fy(y) + C] dy dx.gc3** which plots f(x, y) and calculates the volume between the surface of f and the region R.

The example used in this file is f(x, y) = - 2 sin²(x) cos(y) + 2 with x, y out of [-2, 2].

The xy plane R= [-2, 2] x [-2, 2] is violet.



Programming Details:

- After input of f(x, y) the terms $f_x(x)$, $f_y(x)$, and the constant C must be separated manually, resulting in:

  f(x,y) = - 2 sin(x)^2*cos(y) + 2

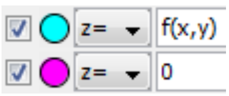
  $f_x(x)$ = - 2 sin(x)^2

  $f_y(y)$ = cos(y)

  C = 2

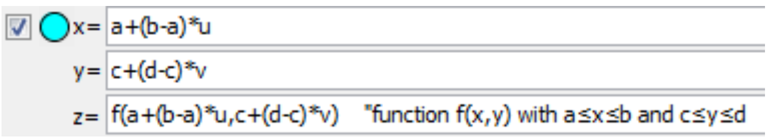
- For setting region R you can use sliders 'a' and 'b' for x, and sliders 'c' and 'd' for y.

- For plotting the function f and region R you may simply use function type '3D Regular'

  The disadvantage of this is that you have to adjust the plot range for f and R as soon as you change the borders of the intervals [a, b] and [c, d].

  An alternative, avoiding manual adjusting of the ranges for f and R according to integral borders, is to use parametric functions for f and R; the transformations

  with v = 0 ... 1 will create the desired intervals [a, b] and [c, d] and will plot the function above the desired region.

  For plotting R the same transformations for x and y are used with z = 0.

- The result is displayed in line 'Integralx (a,b)'.

- **Important**: Make sure to input the terms for fx, fy, and C properly:
  If either f_x or f_y is missing in f(x, y), f_x resp. f_y has to be set to 1.

Further examples for Type 1

- Solve the double integral

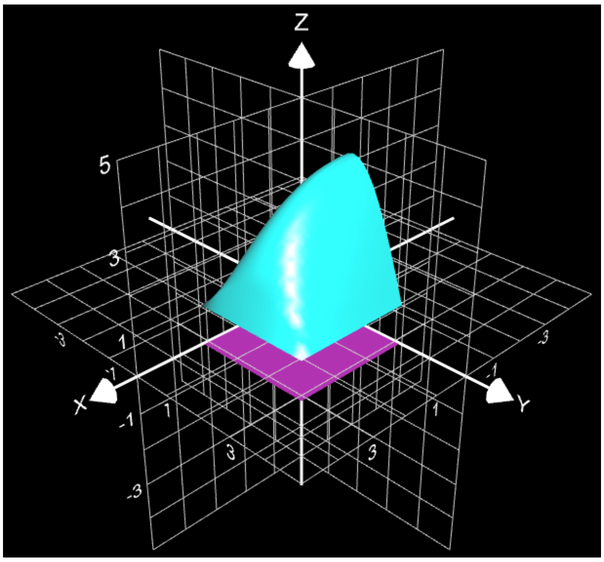
$$\int_0^{\pi}\int_0^{\pi}\left[\sqrt{\pi^2 - x^2}\,\sin(y) + 1\right]\,dy\,dx$$

Exact result::

$$\frac{\pi^2}{2}(2+\pi)$$

Absolute error (n=980): 2.61 E-5

File: **DInt [sqrt(pi^2-x^2)sin(y)+1] dy dx.gc3**



- Solve the double integral

$$\int_{-2}^{2}\int_{-2}^{2}\left[-3e^{-x^2-y^2} + 1\right]\,dy\,dx$$

Make sure to set f_x and f_y properly:

$$-3e^{-x^2-y^2} + 1 = -3e^{-x^2}\,e^{-y^2} + 1$$

$$f_x = -3e^{-x^2}$$

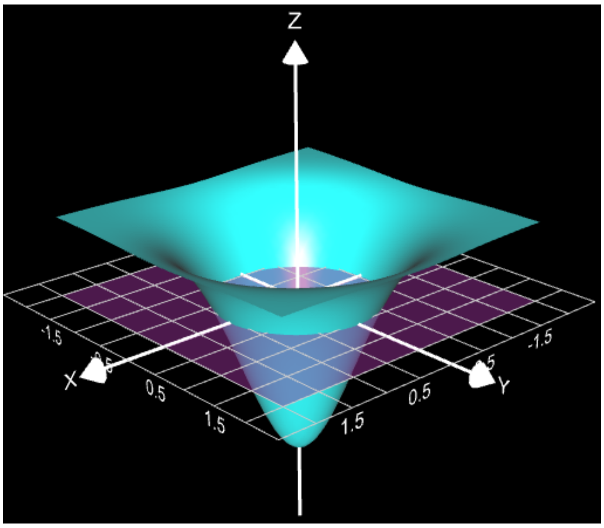
$$f_y = e^{-y^2}$$

C = 1

Exact result: 16 - 3 π erf (2)²

Absolute error (n=980): 1.5 E-12.

File: **DInt [-3e^(-x^2-y^2)+1] dy dx.gc3**



Type 2: f(x, y) = $f_x(x) \pm f_y(y)$ + C

With this type for the terms f_x and f_y the above mentioned also applies with the exception that the terms are not multiplied but added / subtracted and that $f_x(x)$ resp. $f_y(y)$ must be set to zero if either f_x or f_y is missing in f(x, y).

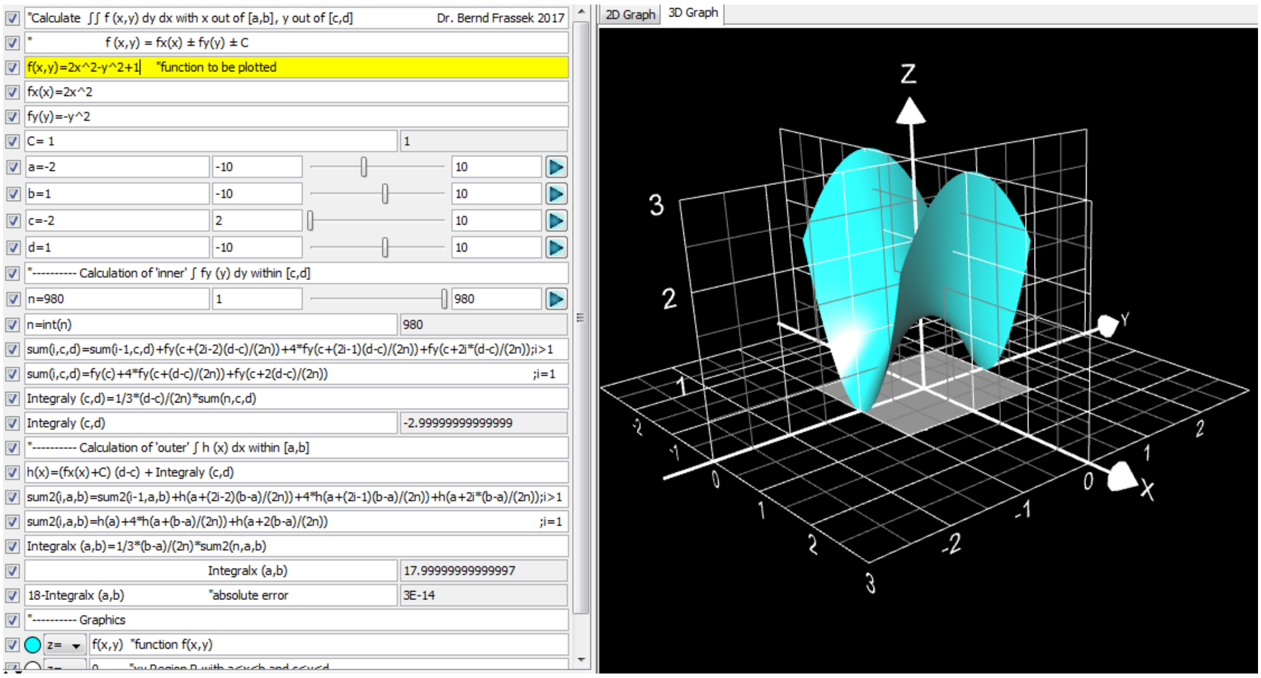
The double integral for this type of f(x, y) can be integrated in the following way:

$$\int_a^b\int_c^d f(x,y)\,dy\,dx = \int_a^b\int_c^d \left[f_x(x) \pm f_y(y) + C\right]dy\,dx$$

$$= \int_a^b\left[\int_c^d \left[f_x(x) + C\right]dy \pm \int_c^d f_y(y)\,dy\right]dx = \int_a^b\left[\left(f_x(x)+C\right)(d-c) \pm \int_c^d f_y(y)\,dy\right]dx = \int_a^b h(x)\,dx$$

The above formulas are implemented in the file **DInt [fx(x) + fy(y) + C] dy dx.gc3** which plots f(x, y) and calculates the volume between the surface of f and the region R.

The example used in this file is: $f(x, y) = 2x^2 - y^2 + 1$ and the handling is the same as with Type 1.



With the exact result being 18, GC3 calculates the volume (n=980) with an absolute error of 3E-14.

Further Example for Type 2

- Solve the double integral

$$\int_{-5}^{5}\int_{-5}^{5} \left[\sin(0.2\,x^2) + \cos\,(0.2\,x^2) + 2\,\right] dy\,dx$$

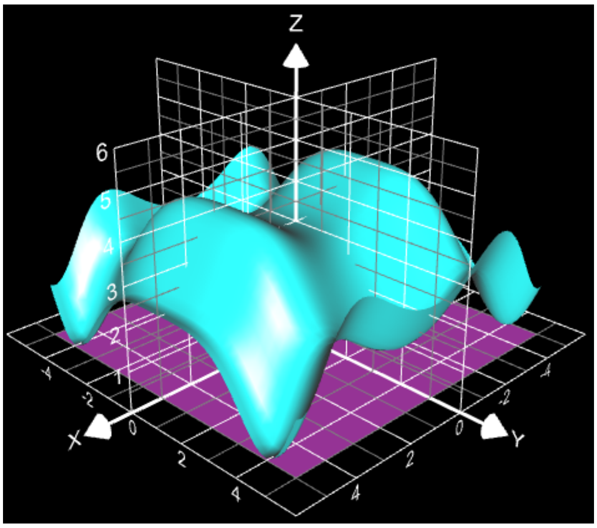
Exact result:

$$10\sqrt{10\pi}\left(C\left(\sqrt{\frac{10}{\pi}}\right) + S\left(\sqrt{\frac{10}{\pi}}\right)\right) + 200$$

(see chapter 3 for Fresnel Integral S (x))

Absolute error (n = 980): 6.26 E-10

File: **DInt [sin(0.2x²)+cos(0.2y²)+2] dy dx.gc3**



5. Area between two Functions

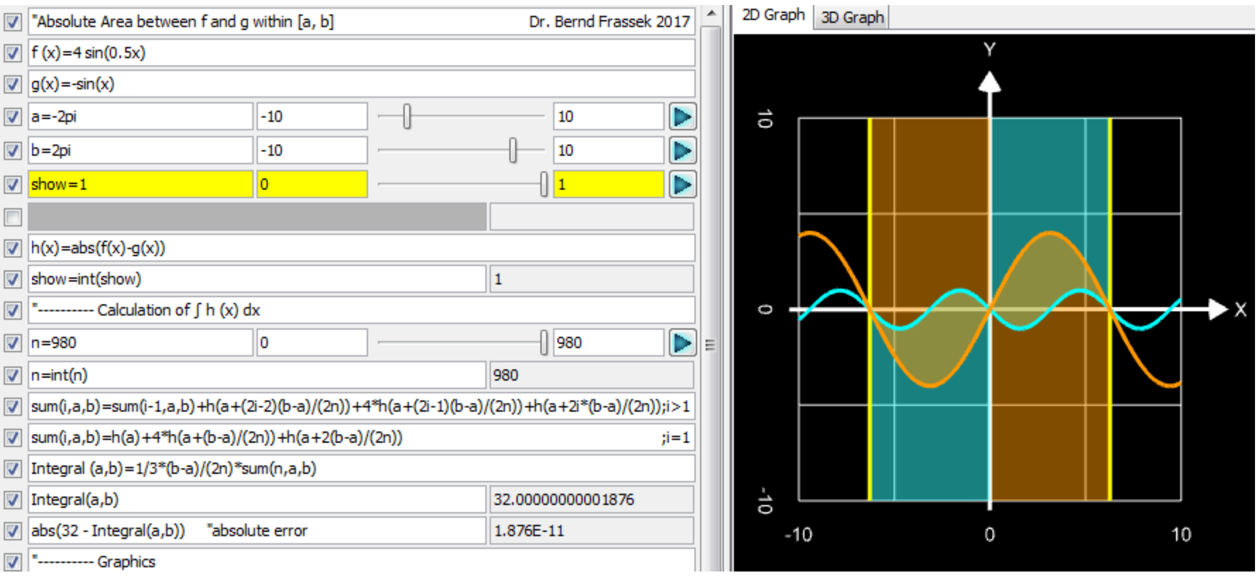
Another task for integration is the calculation of the area between two functions f and g within an interval [a, b] for x.

Especially when the functions intersect each other within [a, b] this can be done 'elegantly' by means of numerical integration without having to determine the points of intersections and taking care of which part of one function is above or below the other function:

$$A = \int_a^b |\, f(x) - g(x) \,| dx$$

The file **Area between 2 Functions.gc3** shows an example for f(x) = 4 sin(0.5x) (orange line) and g(x) = - sin(x) (blue line).

Again, the calculation is done numerically as above but the function h = | f(x)-g(x) | is integrated.

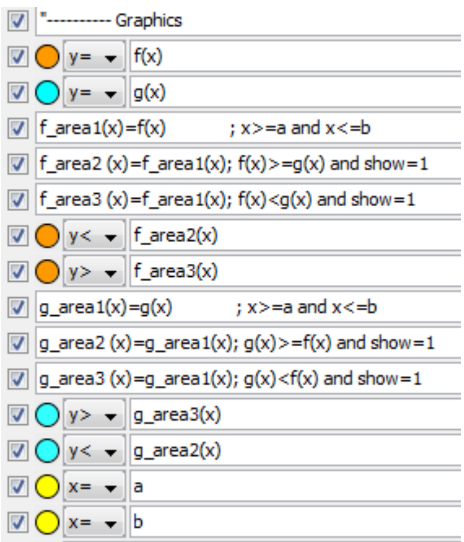


The absolute error is 1.876 E-11.

Programming Details

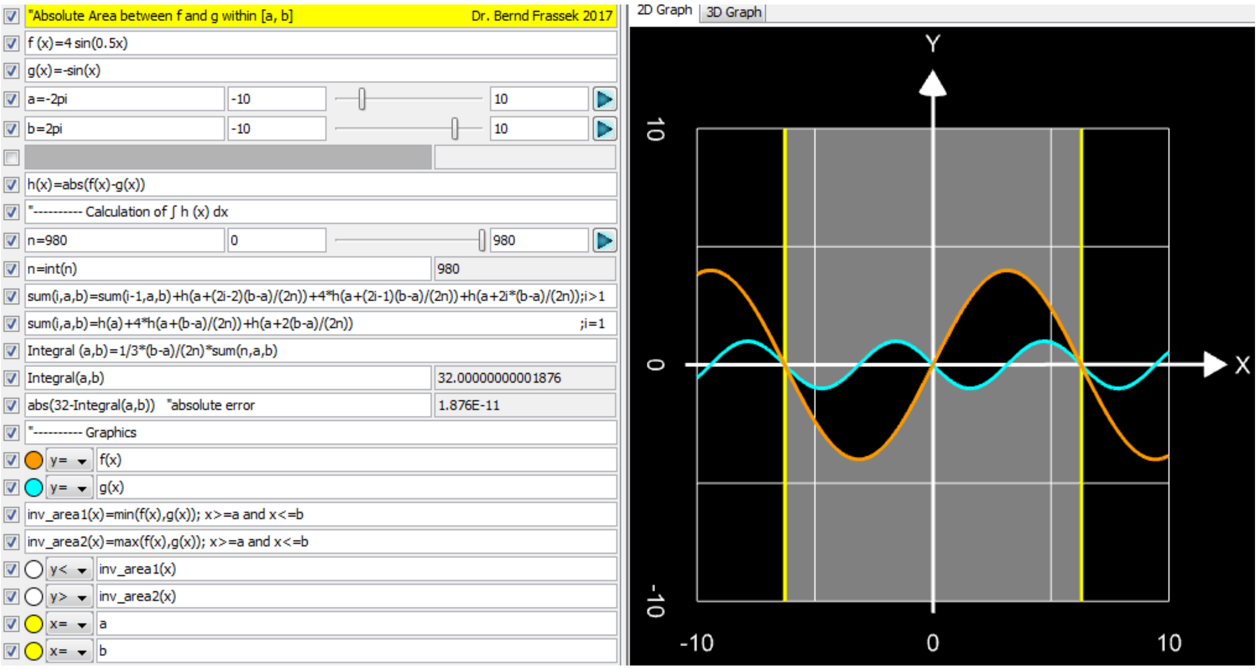
- The coloring of the area between f and g is a little 'tricky' and effectuated by using the ability of GC3 to plot inequalities 'y<' and 'y>' with appropriate case differentiations.

  The area A is a mix of the colors for the inequalities (orange and blue).



- With the variable 'show' you can quickly turn the coloring of the area on or off.
- Another (easier) approach for coloring the area A is done in the file **Area between 2 Functions – inv area.gc3**.

  With this implementation the area is 'empty' while the complete xy area of the interval [a, b] is white.



Quick References:

[1] Integral: https://en.wikipedia.org/wiki/Integral

[2] Antiderivatives: https://en.wikipedia.org/wiki/Lists_of_integrals

[3] Numerical Integration: https://en.wikipedia.org/wiki/Simpson%27s_rule

[4] Numerical Int.: http://www2.math.umd.edu/~dlevy/classes/amsc466/lecture-notes/integration-chap.pdf

[5] Fresnel Integral: https://en.wikipedia.org/wiki/Fresnel_integral

[6] Normal Distribution: https://en.wikipedia.org/wiki/Normal_distribution

[7] CDF: https://en.wikipedia.org/wiki/Cumulative_distribution_function

[8] Error Function: https://en.wikipedia.org/wiki/Error_function

[9] Double (Surface) Integral: https://en.wikipedia.org/wiki/Surface_integral